知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum
E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

基于知识图谱的国内数据素养研究述评

李姣姣[1] 杨志和[2]
1. 中国信息通信研究院 北京 100080
2. 黑龙江大学信息管理学院 哈尔滨 150081

摘要：［目的 / 意义］了解国内数据素养研究现状，梳理数据素养研究主题内容以及发展趋势。［方法 / 过程］以共词分析方法为理论基础，基于聚类分析法和战略坐标分析法，构建数据素养研究主题知识图谱及战略坐标图，并以可视化方式展现。［结果 / 结论］数据素养研究处于探索期，研究内容呈现多元化，大致分为 8 个研究主题，不同的研究主题之间相互关联、相互促进，未来的研究趋势可能向图书馆数据素养教育研究和新一代图书馆馆员的数据素养研究方向跃进。

关键词：数据素养 知识图谱 共词分析 战略坐标分析
分类号：G252

引用格式：李姣姣，杨志和．基于知识图谱的国内数据素养研究述评 [J/OL]. 知识管理论坛，2022, 7(1): 153-163[引用日期]. http://www.kmf.ac.cn/p/283/.

1 引言

2007 年 E. Stevenson 等正式提出了数据素养的概念，将“找到、评价以及合理有效利用信息”作为数据素养的核心要点[1]。黄如花等解析数据素养内涵不再拘囿于数据采集、表示和描述等数据能力层面，还将数据意识、数据伦理纳入数据素养的定义标准[2]。孟祥保等强调数据素养要求科研工作者对数据自身所具有的判断、分析以及转换技能[3]。近年来，数据素养不断成为图书情报界研究的热点话题，为了解国内数据素养研究现状、梳理数据素养研究主题内容以及发展趋势，本文以共词分析方法为理论基础，采用聚类分析算法和战略坐标分析法，构建数据素养研究主题知识图谱及战略坐标图，并以可视化方式展现，为业界数据素养教育未来的发展提供些许参考。

2 研究思路及方法

2.1 研究思路

基于中国知网（CNKI）数据库，以中文社会科学引文索引（CSSCI）和《中文核心期刊目录总览》收录的期刊作为研究数据，首先基于共词分析理论，借助智能局部移动算法（Smart Local Moving algorithm，SLM）进行主题聚类，以识别数据素养研究主题。其次通过向心度

作者简介：李姣姣，工程师，硕士，E-mail: 18646646027@163.com；杨志和，副教授，博士，硕士生导师。

收稿日期：2021-11-29 发表日期：2022-03-24 本文责任编辑：刘远颖

(Centrality)和密度(Density)构建战略坐标图(Strategic Diagram),分析各主题的发展阶段及关系,进一步阐释数据素养研究的现状以及未来的发展趋势。具体包括以下几个步骤(见图 1):①获取文献数据集,并抽取关键词;②进行数据处理,合并同类型关键词,去掉不具有实际意义的关键词,获得关键词数据集;③基于步骤②中的关键词数据集构建关键词共现矩阵;④利用智能局部移动算法,实施主题聚类并可视化;⑤根据主题聚类的结果,计算各主题的向心度和密度,构建战略坐标图,分析各主题间的关系;⑥对研究趋势的分析。

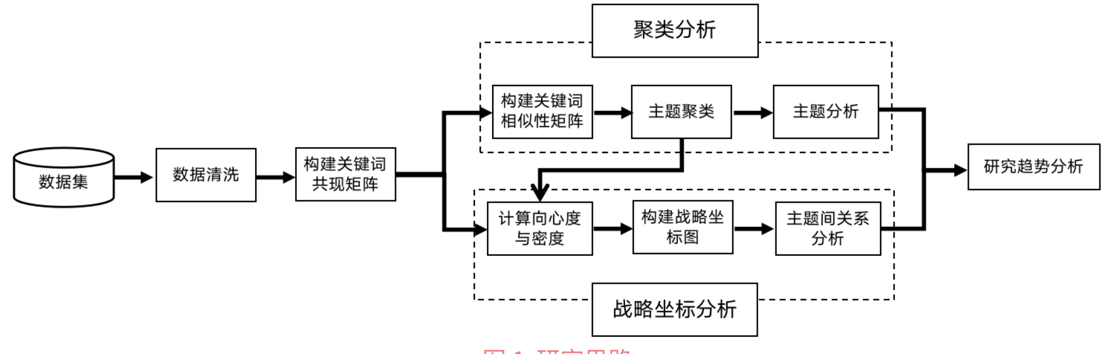


图 1 研究思路

2.2 研究方法

共词分析是内容分析中的一种方法,由 M. Callon 等提出 [4],主要通过词(或词组)的共现反映文本的主要内容 [5]。聚类分析由 H. E. Driver 和 A. L. Kroeber 在 1932 年首次引入人类学的研究中 [6],是一种无监督学习算法,主要依据"物以类聚"的原理,将具有相似性的研究对象聚集在一起。本文将基于智能局部移动算法,借助于 VOSviewer 1.6.16 软件实现主题聚类 [7]。

战略坐标分析由 J. Law 等提出 [8],用于分析特定领域中研究主题之间的关系及主题的发展阶段。战略坐标图是战略坐标分析的主要呈现方式,是一个二维坐标图,横坐标用向心度(Centrality)表示,纵坐标用密度(Density)表示。对于给定主题类簇,向心度表示该主题类簇与其他主题类簇之间联系的强弱,密度表示该主题类簇内部词与词之间联系的强弱 [9],向心度与密度的计算方式 [10] 如下:

$$D(k)=\frac{\sum_{i=0}^{n_k}\sum_{j=0}^{n_k}c_{ij}}{n_k-1}(i\neq j)\qquad \text{公式(1)}$$

$$C(k)=\frac{\sum_{i=0}^{n_k}\sum_{j=0}^{N-n_k}c_{ij}}{N-n_k}\qquad \text{公式(2)}$$

其中,$D(k)$ 是主题类簇 k 的密度,$C(k)$ 是主题类簇 k 的向心度,n_k 是主题类簇 k 的关键词个数,N 是共词矩阵中所有关键词数目,c_{ij} 表示关键词 i 和关键词 j 的共现频次。

根据每个类簇的向心度与密度值构建战略坐标图,以向心度为横坐标,密度为纵坐标,见图 2。第一象限中的主题类簇有较强的密度和向心度,表明该研究主题与其他主题之间联系紧密且该主题内部的关联度均较强,是该领域较为核心且成熟的研究主题;第二象限中的主题密度较高、向心度较低,表明主题与主题之间的关联度较弱,但主题内部的关联度较强,是该领域较为边缘但研究成熟的主题;第三象限中的主题密度和向心度均较低,表明主题与主题之间的关联度较弱,但主题内部的关联度也较弱,是该领域较为边缘且研究不成熟的主题,可能是衰退型的研究主题,也可能是新兴研究主题;第四象限中的主题密度较低、向心度较高,表明主题内部的关联度较弱,但主题

ChinaXiv合作期刊
知识管理论坛，2022（2）:153-163
DOI: 10.13266/j.issn.2095-5472.2022.013

与主题之间的关联度较强，是该领域较为核心但研究不成熟的主题，具有较大的发展潜力[11-12]。



图 2 战略坐标图

3 数据获取与分析

3.1 数据来源及处理

以“数据素养”为主题，基于中国知网（CNKI），选择高级检索方式，选取 CSSCI（中国社会科学引文索引）和《中文核心期刊目录总览》收录的期刊作为研究数据，时间范围不限，检索时间为 2021 年 9 月 2 日，获得检索结果 401 篇，剔除新闻报道、征文通知及不相关的文献，最终获得 395 篇文献。

3.2 数据分析

3.2.1 发文量与引文分析

学术文献的时间序列分析是指特定领域文献数量的逐年产出情况，用于揭示科学发展的规律、判断科学发展的态势。首先，对数据素养相关研究的发文量及总被引频次进行时间序列的分析。从图 3 可以看出，国内数据素养的研究大致从 2013 年开始；2014 年至 2016 年，数据素养相关主题的发文量及总被引频次迅速增加，2016 年总被引频次达到 1 227 次，总发文量为 53 篇；2016 年之后发文量呈现缓慢增长，于 2019 年之后发文量又呈现显著增长态势。总之，在数据爆发式增长环境下，数据素养的研究处于发展阶段。

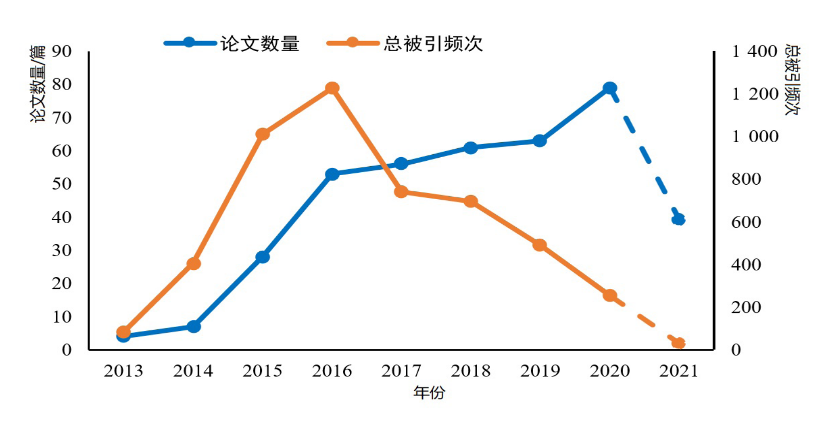


图 3 国内数据素养文献量 / 总被引频次随时间变化趋势

3.2.2 数据素养研究主题分析

抽取频次大于等于 5 的关键词，借助共词分析技术和聚类分析技术绘制国内数据素养聚类分析图。基于 VOSviewer 对国内数据素养研究主题进行聚类分析和可视化（见图 4），数据素养领域含有 8 个研究主题：科研人员数据素养教育研究、数据馆员数据素养教育研究、教师数据素养教育研究、科学数据管理研究、数据素养能力评价研究、图书馆数据素养教育研究、图书情报学科中数据素养教育研究、大数据时代数据素养教育研究。

（1）科研人员数据素养教育研究。该主题主要以数据素养教育、信息素养、素质教育、信息检索、图书馆、数据科学、教学模式、结构方程模型、层次分析等关键词为标识。相关研究者对科研人员数据素养教育的评价进行研究，认为国外高校图书馆针对不同层次和不同领域的研究人员提供个性化的数据素养教育，而国内高校图书馆数据素养教育主要以传统信息素养教育为主，并未系统展开数据素养教育 [13]。梁宇等采用质性研究的方法，基于扎根理论分别从数据维度、博士生自身素养维度、数据环境维度分析评价影响博士生数据素养教育的原因 [14]。惠恭健等通过对比国内外专业数据素养能力模型、公民数据素养能力模型以及学生数据素养能力模型，研究学生数据素养内容及结构模式，并分别从意识、知识与技能、思维及道德规范 4 个维度构建学生数据素养能力评价的指标体系 [15]。

（2）数据馆员数据素养教育研究。该主题包含的核心关键词有大学图书馆、学术图书馆、人才培养、学科馆员、数据馆员、数字人文、数据服务等。研究者们从不同的角度出发，根据不同的案例分析数据馆员的工作内容及服务方式，顾立平等通过对学术图书馆中数据馆员招聘信息的分析，提炼出数据馆员在科研数据管理服务中的工作内容以及服务方式，并基于科研数据全生命周期提出数据馆员数据素养教育的内容，如数据发现能力、数据使用能力、数据采集能力等 [16]。张璐等在对国外数据馆员培训实践案例进行调研分析的基础上提出我国应更加注重数据素养教育、开发培训资源、加强合作交流等 [17]。李梅通过对爱丁堡大学数据管理及服务分析，阐述数据馆员应具备专业数据能力、与高层次学术研究者的合作能力以及团队合作能力的特征 [18]。也有学者分析探索数据馆员或学科馆员的职业技能或职业发展 [19,21]。

（3）科学数据管理研究。该主题主要包含科学数据、科学数据管理、数据监管、数据管理、数据隐私、数据安全、数据共享、数据管理服务、数据生命周期、科研数据、科研人员等关键词。一方面，相关研究者基于数据素养理论，从数据管理知识、数据安全意识以及数据管理技能的维度探究国内外高校科研数据管理服务现状；另一方面，研究者基于科研生命周期的各个阶段，分析科研人员需要具备的科学数据素养 [22]。此外，也有研究者基于国外科学数据管理机构的实践经验，提出我国社会科学数据管理应该重视数据再利用与共享，制定相关政策，统筹规划各级科学数据管理机构，推进科学数据素养教育 [3]。

（4）数据素养能力评价研究。该主题主要以关键词高校图书馆、教育数据、教育职能、教育评价、教育数据、数据密集型科研、数据分析、数据挖掘、数据质量等为代表。数据素养教育评价的研究主要针对高校图书馆和公共图书馆，研究者们主要基于 H. D. Lasswell 的“5W”模式理论对数据素养教育进行评价 [23]，即 Who（数据素养教育主体）、What（数据素养教育内容）、Whom（数据素养教育对象）、Which channel（数据素养教育方式）、What effect（数据素养教育效果）。朱静等从数据素养教育的目标、对象、内容、模式以及评价的角度构建公共图书馆数据素养教育体系 [24]。张长亮等分别以中国和美国的 3 所高校为例，从数据素养教育的目标、内容、方式以及评价体系的角度分析中美的差异性，并提出我国数据素养教育应针对不同群体的需求进行个性化教

知识管理论坛，2022（2）:153-163
DOI: 10.13266/j.issn.2095-5472.2022.013

育、丰富数据素养培训内容、加强多方面合作等[25]。潘雪等基于“5W”模式构建高校图书馆数据素养教育评价体系[26]。

（5）图书馆数据素养教育研究。该主题包含的主要关键词为科学数据素养、数据素养能力、教育体系、教学决策、指标体系、教师等。图书馆数据素养教育的研究主要分为高校图书馆数据素养教育的研究、公共图书馆数据素养教育的研究以及教育机构的数据素养教育的研究，其中高校图书馆数据素养教育的研究是该主题的重点研究内容。吴爱芝等通过对国内高校图书馆数据素养教育形式的调研，以北京大学图书馆为实例，从数据伦理、数据获取、数据处理、数据保持以及评价与引用的角度提出高校图书馆数据素养教育的内容及体系[27]。张文亮等基于国外数据素养教育的现状，结合我国的实际情况，从数据素养教育的对象、结构及内容，提出了我国高校图书馆数据素养教育的框架[28]。肖希明等以中美两国拥有博士学位授予权的图书情报（LIS）教育机构为例，从课程内容、培养意识、培养力度的角度分析了两国 LIS 机构的数据素养教育现状，并提出我国 LIS 机构数据素养教育应该加强数据意识、构建全方位的数据素养教育体系、加强与不同学科之间的合作[29]。

（6）教师数据素养教育研究。该主题主要用关键词教师数据素养教育、教师专业发展、中小学教师、数据应用、数据思维、数据意识、数据技能、数据驱动教学、评价指标、因子分析、德尔菲法等来诠释。教师数据素养教育的研究包含国外教师数据素养教育的研究、中小学教师数据素养教育的研究、教师数据素养教育的评价以及教师职业发展的研究。李青等通过对国外教师数据素养培训项目进行调研，分析了教师数据素养教育的形式及内容，并提出教师数据素养应该从政府、学校以及教师个人 3 个层面展开[30]。同时，李青等还从知识技能、教学实践以及教学探究 3 个层面探索并构建了教师数据素养能力模型[31]。此外，中小学教师数据素养教育的研究成为该主题的研究重点，包括中小学教师数据素养教育内容[32]、评价体系[33]、发展路径及策略[34]等研究。

（7）大数据时代数据素养教育研究。该主题主要以大数据、人工智能、大数据分析、大数据素养、教育信息化、教育治理、数据新闻、智慧教育、精准教学等关键词为代表。研究者们逐渐开始探索大数据环境下数据素养教育的内容、培养路径以及评价指标体系，惠恭健等通过对已有国内外数据素养能力模型的分析，从数据道德规范、数据意识、数据技能及数据思维的角度构建了智能时代我国学生数据素养能力模型，并提出学生数据素养能力评价的指标体系[15]。许亚锋等提出在人机协同教学环境下，教师应该具备基于人工智能技术的数智素养，并从知识与技能、思维能力以及伦理规范的视角提出数智素养的构成要素、作用机制以及培养方式，为教师数据素养教育的研究提供了新的思路[35]。

（8）图书情报学科中数据素养教育研究。该主题主要以图书情报教育、信息素养教育、政府数据、开放数据、培养路径、应对策略、大数据时代等关键词为代表。研究者们逐渐开始探索图书情报学科中数据素养教育的发展路径。黄如花等基于学界和业界的研究现状阐释了大数据背景下数据素养的起源及内涵，同时提出数据素养教育的内容、主体、对象和形式更加多元化[3]。司莉等对 38 所院校的 138 个图书情报专业的数据素养课程进行调研分析，从数据伦理、数据技能、数据意识 3 个层面对数据素养课程内容进行分析，发现当前图书情报专业关于数据素养课程的设置逐渐趋于系统化，且面向大数据处理技术的数据挖掘类课程也在迅速发展[36]。周林兴等通过对国内高校图书情报专业所开展的数据素养课程进行调研分析，从学科发展及社会发展的需求、数据素养课程体系进程的建设、数据资源和社会资源的角度构建了大数据环境下图书情报专业学生数据素养教育提升的优化策略[37]。

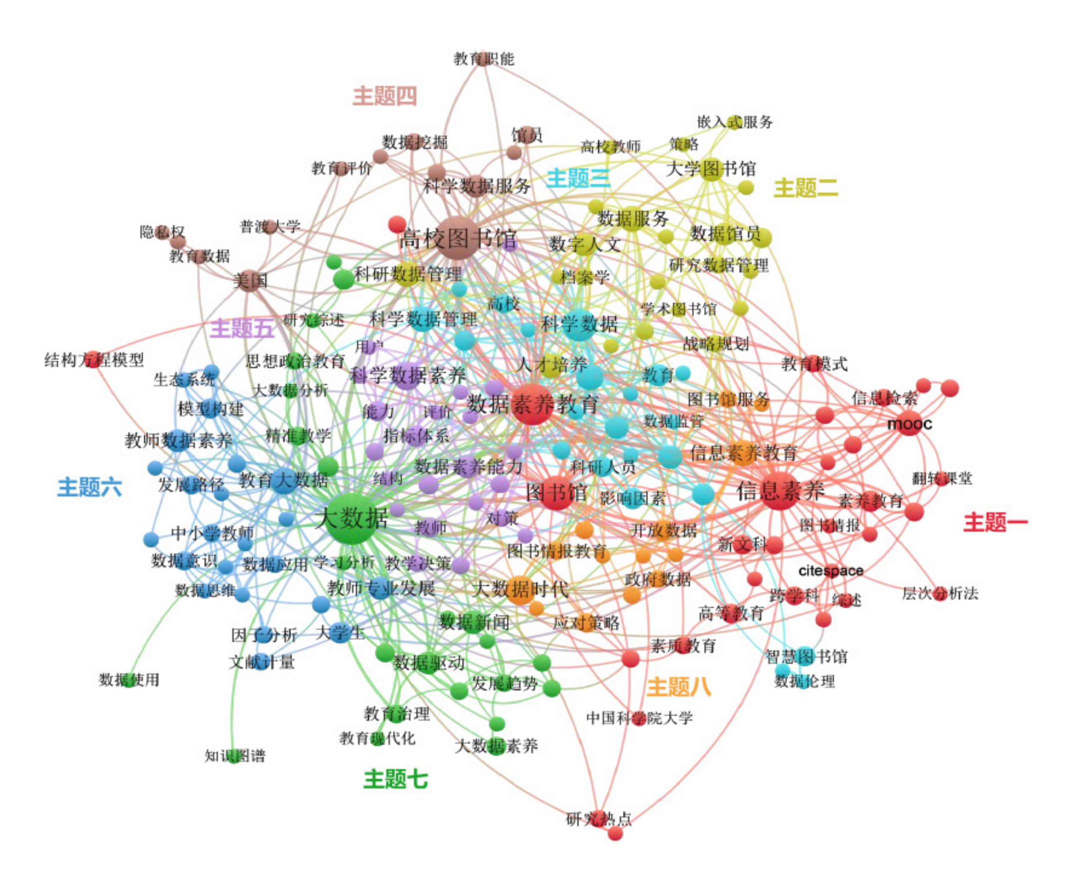


图 4 国内数据素养研究主题

3.2.3 数据素养研究主题的战略坐标分析

首先，基于3.2.2 节中讨论的数据素养主题，根据公式（1）和公式（2）分别计算各主题的向心度和密度（见表 1）。其次，以向心度为 X 轴，密度为 Y 轴，各主题向心度与密度的平均值为原点，绘制战略坐标图（见图 5）。

表 1 各主题向心度和密度值

主题类型	主题内容	向心度	密度
主题一	科研人员数据素养教育	0.895	3.613
主题二	数据馆员数据素养教育	0.368	2.636
主题三	科学数据管理	0.301	3.333
主题四	数据素养能力的评价	0.081	2.385
主题五	图书馆数据素养教育	0.243	2.409
主题六	教师数据素养教育	0.191	2.682
主题七	大数据时代数据素养教育	0.623	3.130
主题八	图情学科中的数据素养教育	0.230	2.231

ChinaXiv合作期刊
知识管理论坛，2022（2）:153-163
DOI: 10.13266/j.issn.2095-5472.2022.013

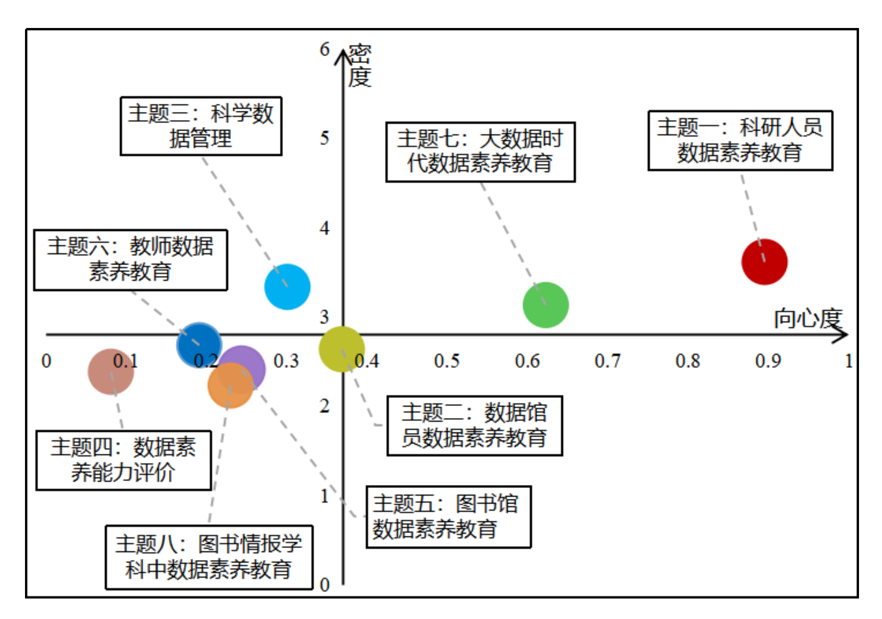


图 5 数据素养主题的战略坐标

向心度代表数据素养研究主题间关系的强弱，即核心度，向心度越大，表明该研究主题与其他研究主题间的关系越紧密，则越是数据素养领域的核心研究主题，反之则是边缘性研究主题。密度代表数据素养研究主题内部的紧密程度，即成熟度，密度越大，表明该主题内部的联系越紧密，主题成熟度越大，反之亦然。

根据数据素养主题的战略坐标分析图（见图 5），从全局角度分析，科研人员数据素养教育（主题一）和大数据时代数据素养教育研究主题（主题七）的向心度值较大，表明这两个研究主题与其他研究主题有较为密切的联系，在数据素养领域中占据着核心地位，而且也是被广泛关注的研究主题；各研究主题的密度值差距不大，表明各研究主题的研究依然处于探索阶段。从局部角度分析，在数据素养领域，相较于其他研究主题，主题一和主题七位于第一象限，属于研究成熟且较为核心的研究主题，这类主题与其他主题的关系较为紧密，且主题内部的连接较为紧密，初步形成研究体系，可以认为是目前数据素养领域的研究热点；主题三位于第二象限，属于研究成熟但边缘化的主题，该主题的研究较为成熟，内部连接较为紧凑，已形成一定的研究规模和体系，但该类主题与其他研究主题的关系较弱一些。主题二、主题四、主题五、主题六、主题八位于第三象限，这类主题属于研究不成熟且边缘化的研究主题，该类主题与其他主题间的关系较为稀疏，且主题内部的结构较为松散，并未形成成熟稳定的研究体系，但随着数字技术的不断发展和应用，这类研究主题未来可能会发展成为数据素养领域新兴的研究主题。

4 研究结论

根据前文的数据素养主题聚类及战略坐标分析结果可知，国内数据素养的未来研究趋势将更倾向于数据素养教育研究。其中，主要研究内容可以分为图 5 中的八大研究主题，并且根据数据素养研究主题间的内容关系可知，八大研究主题可以归纳为 5 类，即主体、客体、目标、效果及环境，其构成了未来数据素养教育的主要核心要素，其具体内容如下：

4.1 图书馆和图书情报学科是推动数据素养教育发展的主体

在过去的信息时代发展主题下，图书情报相关学科就是我国信息素养培养的主体。虽然

在新一代信息技术发展下当前我国进入了全新的数字经济时代，但从“数据—信息—知识—情报”的关联逻辑上看，数据素养的内涵和外延均比过去的信息素养的内涵及外延要宽泛很多。依据 2021 年 11 月 5 日中央网络安全和信息化委员会发布的《提升全民数字素养与技能行动纲要》可知，数据素养内涵、行动方向实际上是对过去信息时代中全民信息素养培养的更深层次内涵挖掘和落实，这使得从前从事图书馆和图书情报学科中信息素养教育科研的更多群体进一步将重心聚焦于推动和完善社会各产业数字化转型升级中的数据素养培养体系构建工作。当前及未来的社会各产业数字化转型升级无疑需要最大范围的实现全民数据素养水平提升，这自然赋予了图书情报学界同仁们需要加大力度充分围绕着国家数字化转型的各个方面开展专业化数据素养教育及发展的主体使命。

4.2 数据馆员、科研人员、教师等是数据素养教育的主要客体

事实上，图书馆在发挥社会服务的载体功能时，其又是数据素养教育的主要客体。数据素养教育是各类图书馆的重要职能之一，特别是图书馆中的数据馆员则是未来数字时代图书馆发挥社会服务的重要载体；当然，从“数据—信息”服务的顶层逻辑来看，社会中所有从事数字化转型的科研人员、教师及推动数字化转型的一切人员均构成了数据素养教育的主要客体。2021 年 12 月 27 日中央网络安全和信息化委员会印发了《“十四五”国家信息化规划》，其中将“全民数字素养与技能提升行动”作为十大优先内容之一，这为未来我国数据馆员、科研人员及教师等从事数据素养教育实践与研究的高质量发展指明了行动方向。

4.3 实现科研数据的高效管理是数据素养教育拟达到的主要目标

对于个体来讲，接受数据素养教育的主要目标是提升数据挖掘、分析及应用技能，但从国家战略的高度来看，数据素养教育的主要目标是为了让数据要素在其生产、流动和开发应用的各个环节中均按特定的科学逻辑产生协同与创新效应。因此，对于整个社会来讲，实现科研数据的高效管理自然是当前及未来我国社会各产业数字化转型先期的重要保障手段，同时也是全民数据素养提升所追求的主要目标。

4.4 数据素养教育评价是为了展现数据素养教育的最终效果

和以往信息时代中的信息素养教育成效评价一样，在当前以数据为核心要素的数字化、智能化时代背景中，数据素养教育成效的评价依然非常重要，这关系到数据产业相关人才的高质量培养。此外，在当前的数字人文发展背景下，数据素养教育评价也关系着各学科教育成效的判定。当然，随着大数据、人工智能等相关技术、理论的发展，数据素养教育评价的方法与以往的信息素养教育评价不同，当前及未来的数据素养教育评价需要通过充分挖掘数据要素主体、客体及其作用方式等全要素数据为评价对象，并应用数字化、智能化的科学评价手段及分析方法进行数据素养教育评价，从而对数据素养教育的效果进行实时跟踪与科学反馈，最终实现高质量的数据素养教育评价及其保障体系构建。

4.5 数智化时代背景是数据素养教育所处的环境

在大数据、物联网、云计算、人工智能等全新一代信息技术的发展下，社会各产业正面临着更高水平的数字化、智能化转型升级，进而使数据要素成为未来社会生产中的核心战略资源。在 2021 年 3 月全国两会发布的《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》中专篇指明了要“加快数字化发展，建设数字中国”，其中明确指出要“激活数据要素潜能”。对于数据要素的挖掘及开发利用来说，一旦脱离其所关联的数字化、智能化环境，就会失去其本身所具有的价值。可以说，数据素养的培养需要所有参与方充分把握数据生产、清洗、存储、更新、协同和失效的全周期。因此，在社会各产业广泛

知识管理论坛，2022（2）:153-163
DOI: 10.13266/j.issn.2095-5472.2022.013

进行数字化、智能化转型升级的时代背景中，数据素养教育的主体和客体均需要立足于特定数智化环境开展数据素养的培养和提升工作。

综合来讲，图书馆及图书情报学科作为数据素养教育的主要阵地，也是数据素养教育的主体，促进数据馆员、科研人员、教师等数据素养的提升，科研人员等数据素养的提升有助于实现科研数据的管理。同时，通过对主体及客体数据素养教育的评价可以展现数据素养教育的效果，反向促进主体及客体数据素养的提升；大数据、物联网、云计算、人工智能等新兴信息技术发展下，数字化、智能化升级转型所形成的数智化环境对科研数据管理提出了更高的要求，也推动着图书馆、图书情报学科数据素养教育和数据馆员、科研人员、教师等数据素养教育的提升。因此，数据素养各研究主题之间相互依存，相互促进，共同发展。上述各内容的关键要素相互作用关系如图 6 所示：

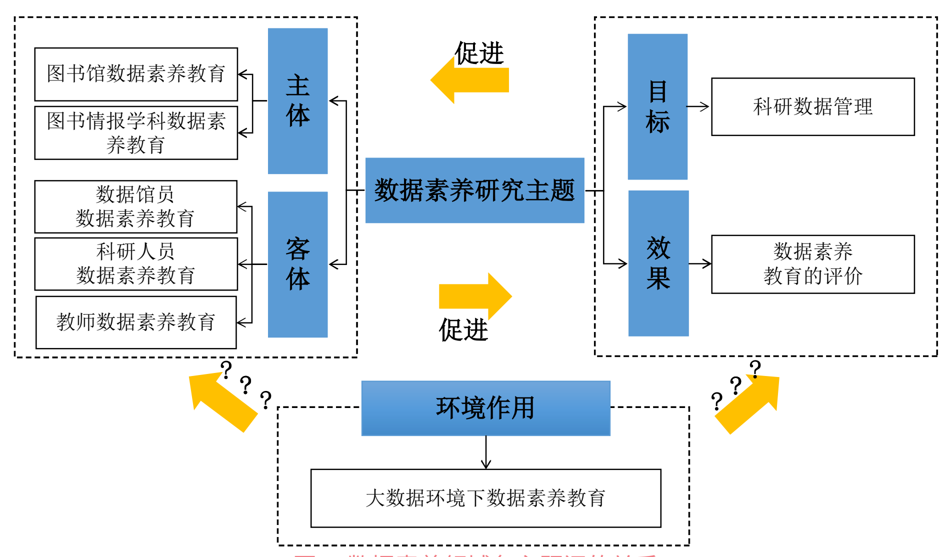


图 6 数据素养领域各主题间的关系

根据数据素养主题战略坐标图（见图 5），“数据馆员数据素养”研究主题的向心度最大（0.368），“图书馆数据素养教育”研究主题的向心度次之（0.243），因此，图 6 数据素养教育研究主体和客体中的图书馆数据素养教育研究和数据馆员数据素养研究将会成为数据素养领域的未来的研究趋势。当前我国数据素养教育的研究依然处于发展的初期，并没有形成完善的数据素养教育理论以及成熟的数据素养教育体系。随着新一代信息技术的不断发展，对公民数据素养的要求会越来越高，图书馆作为数据素养教育的主要阵地之一，有必要构建完善的数据素养教育体系、研究其发展路径，并辅以实践。当然，图书馆数据素养教育的不断发展也将引起图书馆馆员服务模式的变革，开始从原有的学科服务、信息服务转向数据服务，新一代图书馆员（即数据馆员）必将不断兴起，相关的服务体系、岗位职责及数据技能等也会成为主要研究内容。

5 总结

“数智化”趋势下，大数据、人工智能、区块链、云计算等技术的加速创新，推动了数字经济的不断发展。数据作为一种新的生产要素，为数字经济的发展注入了新动能，也对社会的发展产生深刻的影响。数据素养作为数字经济社会发展的基石，有助于人类快速适应新业态、新模式，成为每个人必备的素养，也成为了学术界广泛讨论的焦点。研究发现，数据素养的研究内容更加多元化，数据素养教育的



对象、方式、内容呈现出多样化的发展模式。从数据素养教育对象维度分析，数据素养教育对象已经不囿于传统的图书馆员或者情报分析人员，而是扩大到全民范畴，如教师数据素养教育、科研人员数据素养教育、图书馆员数据素养教育。从数据素养教育方式维度分析，数据素养教育呈现多途径、多渠道发展，后疫情时代，数据素养教育产生了线上线下相结合方式，同时，也在逐渐形成跨学科、跨地域的合作模式。从数据素养教育的内容维度分析，数据素养教育的内容研究更加丰富，除了数据的检索、获取、整理的技能教育培训，数据伦理、数据思维、数据处理、数据挖掘、数据共享、数据保存等内容的研究也成为重点。此外，数据素养教育的内容更加符合实际生产生活。总体分析，数据素养的研究需要与时俱进，数据素养教育的研究更需要满足时代发展的要求，在提升全民数据素养的同时，也需要继续完善数据素养教育的评价体系，推动数据素养的研究。

参考文献：

[1] STEPHENSON E, CARAVELLO P S.Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project[J]. Reference services review, 2007, 35(4): 525-540.

[2] 黄如花 , 李白杨 . 数据素养教育 : 大数据时代信息素养教育的拓展 [J]. 图书情报知识 , 2016(1): 21-29.

[3] 孟祥保 , 钱鹏 . 高校社会科学数据管理的国际经验及其借鉴 —— 以 UKDA 和 ICPSR 为例 [J]. 情报资料工作 , 2013(2): 77-80.

[4] CALLON M, COURTIAL J P, TURNER W A, et al. From translations to problematic networks: an introduction to co-word analysis[J]. Social science information, 1983, 22(2): 191-235.

[5] COBO M J, LÓPEZ-HERRERA A G, HERRERA-VIEDMA E, et al. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the fuzzy sets theory field[J]. Journal of informetrics, 2011, 5(1): 146-166.

[6] BAILEY K D. Numerical taxonomy and cluster analysis[J]. Typologies and taxonomies, 1994, 34: 35-65.

[7] ECK N J V, WALTMAN L. How to normalize cooccurrence data? an analysis of some well - known similarity measures[J]. Journal of the American society for information science and technology, 2009, 60(8): 1635-1651.

[8] LAW J, BAUIN S, COURTIAL J P, et al. Policy and the mapping of scientific change: a co-word analysis of research into environmental acidification[J]. Scientometrics, 1988, 14(3-4): 251-264.

[9] CALLON M, COURTIAL J P, LAVILLE F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemsitry[J]. Scientometrics, 1991, 22(1): 155-205.

[10] LEE B, JEONG Y I. Mapping Korea's national R&D domain of robot technology by using the co-word analysis[J]. Scientometrics, 2008, 77(1): 3-19.

[11] GIANNAKOS M, PAPAMITSIOU Z, MARKOPOULOS P, et al. Mapping child–computer interaction research through co-word analysis[J/OL]. International journal of child-computer interaction, 2020, 23: 100165 [2021-11-01]. https://www.sciencedirect.com/science/article/pii/S2212868920300015.

[12] 王康 , 高继平 , 潘云涛 , 等 . 多位态研究主题识别及其演化路径方法研究 [J]. 图书情报工作 , 2021, 65(11): 113-122.

[13] 雷春蓉 , 陈梦 . 国内外高校图书馆数据素养教育比较研究 [J]. 图书馆 , 2021(4):47-51.

[14] 梁宇 , 郑易平 . 高校博士生数据素养的影响因素及应对策略 [J]. 情报理论与实践 , 2021, 44(4):146-152, 137.

[15] 惠恭健 , 曾磊 . 智能时代的数据素养 : 模型构建、指标体系与培养路径 —— 基于国内外模型的比较分析 [J]. 远程教育杂志 , 2021, 39(4): 52-61.

[16] 顾立平 , 张潇月 . 开放科学环境下数据馆员的实践探析 [J]. 图书情报知识 , 2020(2): 60-74, 112.

[17] 张璐 , 盛兴军 . 中外数据馆员培训实践调研与比较研究 [J]. 新世纪图书馆 , 2020(2): 31-35, 47.

[18] 李梅 . 大学图书馆的研究数据服务创新与数据馆员新角色 —— 英国爱丁堡大学范例研究 [J]. 图书与情报 , 2019(3):122-130.

[19] 王宁 , 王震宇 . 大数据驱动下高校图书馆员数据技能分级培育模式研究 [J]. 高校图书馆工作 , 2018, 38(4):10-17, 47.

[20] 陈媛媛 , 柯平 . 大学图书馆数据馆员职业技能发展研究 [J]. 情报资料工作 , 2018(2): 97-102.

DOI: 10.13266/j.issn.2095-5472.2022.013

[21] 许鑫 , 张月 . 数据时代图书馆员角色的转变 —— 数据馆员的兴起 [J]. 图书情报工作 , 2021, 65(7): 60-66.
[22] 李文文 , 周萍 . 基于科研生命周期的科学素养教育研究 [J]. 现代情报 , 2017, 37(1): 156-159.
[23] LASSWELL H D. The structure and function of communication in society[J]. The communication of ideas, 1948, 37(1): 136-139.
[24] 朱静 . 公共图书馆数据素养教育体系探析 [J]. 高校图书馆工作 , 2020, 40(5): 80-84.
[25] 张长亮 , 王晨晓 , 李竟彤 . 大数据时代中美高校数据素养教育比较研究 [J]. 情报理论与实践 , 2019, 42(8): 131-137.
[26] 潘雪 , 陈雅 . 我国高校数据素养教育评价指标体系探析 [J]. 新世纪图书馆 , 2018(7): 31-34.
[27] 吴爱芝 , 王盛 . 高校图书馆数据素养教育体系设计研究 —— 以北京大学图书馆为例 [J]. 大学图书馆学报 , 2020, 38(6): 96-103.
[28] 张文亮 , 刘敬仪 . 我国高校图书馆数据素养教育体系框架初探 [J]. 高校图书馆工作 , 2017, 37(4): 80-84.
[29] 肖希明 , 倪萍 . 中美 LIS 教育中数据素养教育的调查与比较分析 [J]. 图书与情报 , 2018(1): 20-27.
[30] 李青 , 任一姝 . 国外教师数据素养教育研究与实践现状述评 [J]. 电化教育研究 , 2016, 37(5): 120-128.
[31] 李青 , 任一姝 . 教师数据素养能力模型及发展策略研究 [J]. 开放教育研究 , 2016, 22(6): 65-73.
[32] 李新 , 杨现民 . 中小学教师数据素养培训课程设计与实践研究 [J]. 中国电化教育 , 2020(5): 111-119, 134.
[33] 林秀清 , 杨现民 , 李怡斐 . 中小学教师数据素养评价指标体系构建 [J]. 中国远程教育 , 2020(2): 49-56, 75, 77.
[34] 林秀清 , 杨现民 , 李新 . 中小学教师数据素养的发展路径与培养策略 [J]. 现代教育技术 , 2020, 30(1): 59-65.
[35] 许亚锋 , 彭鲜 , 曹玥 , 等 . 人机协同视域下教师数智素养之内涵、功能与发展 [J]. 远程教育杂志 , 2020, 38(6): 13-21.
[36] 司莉 , 姚瑞妃 . 图书情报专业研究生数据素养课程设置及特征分析 —— 基于 iSchool 联盟院校的调查 [J]. 图书与情报 , 2018(1): 28-36, 101.
[37] 周林兴 , 周丽 , 艾文华 . 大数据背景下图情档专业研究生数据素养教育探析 [J]. 图书情报工作 , 2019, 63(19): 57-65.

作者贡献说明：
李姣姣：进行数据处理与分析，撰写与修改论文；
杨志和：撰写与修改论文。

A Review of Data Literacy Research Based on Knowledge Map in China

Li Jiaojiao[1] Yang Zhihe[2]
[1]China Academy of Information and Communications Technology, Beijing 100080
[2]School of Information Management, Heilongjiang University, Harbin 150081

Abstract: [Purpose/Significance] In order to understand the current status of data literacy research in China, sort out the main topics and development trends of data literacy research. **[Method/Process]** Taking the co-word analysis method as the theoretical basis, based on cluster analysis method and strategic coordinate analysis method, this paper constructed the knowledge map of data literacy research topics and strategic coordinate map, and visualized them. **[Result/Conclusion]** It is found that the research on data literacy is in the exploratory period, and the research content is diversified, which is roughly divided into eight topics. Different research topics are interrelated and promote each other. The future research trend may leap forward to the research on library data literacy education and the research on data literacy of a new generation of librarians.

Keywords: data literacy knowledge map co-word analysis strategic coordinate analysis